MONOTONE BOOLEAN FUNCTIONS CAPTURE THEIR PRIMES

JEAN BOURGAIN

Abstract. It is shown that monotone Boolean functions on the Boolean cube capture the expected number of primes, under the usual identification by binary expansion. This answers a question posed by G. Kalai.

1. Introduction

Identifying the interval $\{1, 2, \dots, 2^n - 1\}$ with Boolean cube $\{0, 1\}^n$ by binary expansion, we prove the following

Theorem 1. *Let f be a monotone Boolean function on $\{0,1\}^n \cap [x_0 = 1]$ and assume $\mathbb{E}[f] > c > 0$, with c some constant taken independent of n.*[1] *Denote Λ the Von Mangold function. Then*

$$\sum_{0<x<2^n} \Lambda(x) f(x) > \big(1 - o(1)\big) 2^n \mathbb{E}[f]. \tag{1.1}$$

This answers an issue brought forward by G. Kalai as part of a collection of problems related to circuit complexity, digital aspects and the Fourier-Walsh spectrum of the Moebius and Von Mangoldt functions. See also [Gr].

In the special case when f is the majority function, we proved in [B3] de-correlation from Λ, which is a stronger statement since it gives a prime number theorem. What is lost in the present context of a monotone Boolean function is the invariance under the permutation group of $\{0, 1, \dots, n-1\}$, an essential ingredient in [B3]. In [B3], we replaced Λ by its symmetrization Λ_s under the permutation group, noting that $\langle f, \Lambda\rangle = \langle f, \Lambda_s\rangle$. This distribution Λ_s turns out to be much better behaved than Λ, in the sense that the high order contribution of the Fourier-Walsh (F-W) spectrum may be handled by L^2-estimates.

The research was partially supported by NSF grants DMS-0808042 and DMS-0835373.

[1] A more precise quantitative version appears in (3.19) below.

At this point, we also invoke the tail estimate

$$\sum_{|S|>k} |\hat{f}(S)|^2 \lesssim k^{-\frac{1}{2}} \tag{1.2}$$

when f is the majority function. For general monotone Boolean functions, (1.2) has a counterpart due to Bshouty and Tamon [B-T].

Proposition 1. *Let f be a monotone Boolean function on $\{0,1\}^n$. Then for $K>1$, we have a tail estimate*

$$\sum_{|S|>K\sqrt{n}} |\hat{f}(S)|^2 \lesssim K^{-1} \tag{1.3}$$

On the other hand, Proposition 2 in the next section establishes a strong bound on $\sum_{\substack{|S|<n_0 \\ S\neq\phi,\{0\}}} |\hat{\Lambda}(S)|^2$ for n_0 as large as $n^{\frac{4}{7}-\varepsilon}$ and hence, one may hope to exploit (1.3). Again a direct L^2-approach fails and Λ needs to be replaced by a friendlier distribution. Keeping the monotonicity of f in mind, we define

$$\tilde{\Lambda} = \sum_{x\in\{0,1\}^n} \Lambda(x) \sum_{j|x_j=1} \delta_{x\backslash\{j\}} \tag{1.4}$$

noting that

$$\langle\tilde{\Lambda}, f\rangle = \sum_x \Lambda(x) \sum_{j|x_j=1} f(x\backslash\{j\}) \leq \sum_x \Lambda(x)\Big(\sum x_j\Big) f(x). \tag{1.5}$$

On the arithmetic side, to each prime

$$p = 2^{j_1} + 2^{j_2} + \cdots + 2^{j_k} \qquad (j_1 = 0)$$

we associate the k integers $x = 2^{j_2}+\cdots+2^{j_k}, 2^{j_1}+2^{j_3}+\cdots+2^{j_k}, \ldots, 2^{j_1}+\cdots+2^{j_{k-1}}$, noting that $f(x)\leq f(p)$.

It turns out that

$$\|\tilde{\Lambda}\|_2 \sim \|\tilde{\Lambda}\|_1 \sim n \tag{1.6}$$

(with (1.6) expressed with normalized measure). Further more, the F-W coefficients $\hat{\tilde{\Lambda}}(S)$ may be retrieved from the $\hat{\Lambda}(S)$, which makes Proposition 2 also applicable to $\tilde{\Lambda}$.

In §2, we prove Proposition 2, a result of independent interest. The argument is in fact quite similar in many aspects to the analysis in [B1] and we only indicate the main points.

In §3, the function $\tilde{\Lambda}$ is analyzed and the scheme described above worked out; inequality (1.1) is established in a more precise form.

As pointed out earlier, (1.1) does not provide an asymptotic formula and it may be unreasonable to expect one in this generality.

Acknowledgement: The author is grateful for various communications with G. Kalai on the subject.

2. On the Fourier-Walsh coefficients of the von Mangoldt function

Let $N = 2^n$ and identify $\{0, 1, \dots, N-1\}$ with $\{0, 1\}^n$ by binary expansion

$$x = \sum_{0 \leq j < n} x_j 2^j \quad \text{with} \quad x_j = 0, 1. \tag{2.1}$$

The Walsh system $\big\{w_S; S \subset \{0, 1, \dots, n-1\}\big\}$ is defined by

$$w_S(x) = \prod_{j \in S} \varepsilon_j \quad \text{with} \quad \varepsilon_j = 1 - 2x_j \in \{1, -1\}. \tag{2.2}$$

If f is a function on $\{0, 1\}^n$, we have

$$f(x) = \sum_S \hat{f}(S) w_S(x) \quad \text{where} \quad \hat{f}(S) = 2^{-n} \sum_{x \in \{0,1\}^n} f(x) w_S(x). \tag{2.3}$$

Considering the restriction of the Von Mangoldt function Λ to $\{1, \dots, N-1\}$ as a function on the Boolean cube $\{0, 1\}^n$, the following estimate holds on its Fourier-Walsh spectrum $\{\hat{\Lambda}(S)\}$.

Proposition 2. *Let $n_0 < n^{4/7} (\log n)^{-2}$. Then for n large enough*

$$\sum_{|S| \leq n_0, S \neq \phi, \{0\}} |\hat{\Lambda}(S)|^2 < e^{-n^{3/7}}. \tag{2.4}$$

The condition $S \neq \{0\}$ of course is parity related.

Note that for application in conjunction with Proposition 1, it is essential that n_0 brakes the $n^{1/2}$-barrier (cf. the related work by B. Green [Gr] about correlation with $AC(0)$ circuits).

A similar statement was established in [B1] for the Moebius function μ (stated with a worse estimate). Despite many similarities, there are significant differences in the analysis and we rather follow the treatment in [B2]. The problem studied in [B2] is the evaluation of

$$\sum_1^N \Lambda(x) f(x) \tag{2.5}$$

with f of the form $f = 1_{[x_{j_1} = \alpha_1, \dots, x_{j_r} = \alpha_r]}$ (i.e. primes with certain prescribed binary digits). Most of the analysis in [B2] is presented in

the context of a general function f however. The following technical statement summarizes §1, 2, 3 from [B2].

Lemma 1. *Assume f a function on $\{1, \dots, N-1\}$ supported on the odd integers and of mean zero. Then*

$$\Big|\sum_1^N \Lambda(x) f(x)\Big| \lesssim$$

$$\frac{N}{\sqrt{B}} n^3 \|\hat{f}\|_1 \tag{2.6}$$

$$+ B\|\hat{f}\|_1 \Big[\max_{\substack{u<N\\ \mathcal{X}\in\mathcal{G}}} |\psi(u, \mathcal{X})|\Big] \tag{2.7}$$

$$+ n^2 \Big[\sum_{q<B, q\, sf} \frac{\kappa(q)}{q}\Big] \|f\|_1 \tag{2.8}$$

$$+ B \exp\big(-cn^{3/5}(\log n)^{-\frac{1}{5}}\big) \|f\|_1 \tag{2.9}$$

$$+ n^5 \Big\{ {\sum}^*_{\mathcal{X}_1\in\mathcal{B}} \frac{q_1}{\phi(q_1)} \Big[\sum_{q_3<B, q_3 sf,\ odd} \frac{\alpha(q_1, q_3)}{q_3}\Big]\Big\} \|f\|_1. \tag{2.10}$$

Here $\hat{f}$ refers to the usual Fourier transform, $\psi(u, \mathcal{X}) = \sum_{x<u} \Lambda(x)\mathcal{X}(x)$ and B is a parameter. The classes $\mathcal{G}$ and $\mathcal{B}$ corresponding to a subdivision of the Dirichlet characters according to the zeros of their corresponding L-function. More specifically, let $T > B, \log T \sim \log B$ be another parameter and denote

$$\eta(\mathcal{X}) = \min\{1-\beta; \rho = \beta + i\gamma, |\gamma| < T \text{ zero of } L(s, \mathcal{X})\}. \tag{2.11}$$

Then

$$\mathcal{G} = \Big\{\mathcal{X}(\text{mod}\, q) \text{ non-principal}; q < B \text{ and } \eta(\mathcal{X}) > C\frac{\log T}{n}\Big\} \tag{2.12}$$

and

$$\mathcal{B} = \big\{\mathcal{X}(\text{mod}\, q) \text{ non-principal}; q < B\big\} \backslash \mathcal{G}. \tag{2.13}$$

In (2.10), $\sum^*$ refers to summation over primitive characters.

Remains to specify $\kappa(q), \alpha(q_1, q_3)$. They are required to satisfy the respective inequalities

$$\Big|\sum_{x<N, q|x} f(x)\Big| \le \frac{\kappa(q)}{q}\|f\|_1 \text{ for } q < B \text{ odd and square-free} \tag{2.14}$$

and

$$\Big|\sum_{x\in J, q_3|x} f(x)\mathcal{X}_1(x)\Big| \le \frac{\alpha(q_1, q_3)}{q_3} \sum_{x\in J} |f(x)| \tag{2.15}$$

for $q_1, q_3 < B, (q_1, q_3) = 1$, q_3 *odd and square-free,* $\mathcal{X}_1 (mod\, q_1)$ *primitive and* $J \subset [1, N]$ *an arbitrary interval of size* $\sim \frac{N}{B}$.

The above the statement is established using the circle method. A few words of explanation about the different contributions. The term (2.6) is the minor arcs contribution, estimated using Vinogradov's inequality. The major arcs are analyzed the usual way, using Dirichlet characters; (2.8), (2.9) account for the contribution of the principal characters.

The contribution of the non-principal characters $\mathcal{X} \in \mathcal{G}$ is expressed by (2.7) and those in $\mathcal{B}$ by (2.10). It is important to note that the savings in (2.8), (2.10) depend on $\kappa(q)$ and $\alpha(q_1, q_3)$ in inequalities (2.14), (2.15).

Following [B2], §6, take

$$\log T \sim n^{4/7} (\log n)^{-3/7} \tag{2.16}$$

which in (2.7) gives an estimate

$$|\psi(u, \mathcal{X})| \lesssim \frac{n^2 N}{T} \text{ for } \mathcal{X} \in \mathcal{G}, u < N. \tag{2.17}$$

In order to deduce Proposition 1 from Lemma 1, the natural choice for f would be the truncation

$$f = \Big[\sum_{\substack{S \subset \{1, \ldots, n-1\} \\ 0 < |S| < n_0}} \hat{\Lambda}(S) w_S \Big] (1 - \varepsilon_0) = \sum \hat{f}(S) w_S. \tag{2.18}$$

Similarly to [B2], we make the following modification of f.

Let $m \in \mathbb{Z}_+$ be another parameter, satisfying

$$\log B < m, n \sim \log B \tag{2.19}$$

and partition $\{0, 1, \ldots, n-1\}$ in intervals J_α of size $\frac{n}{m}$.

Let

$$K_0 = 1 + 10 \frac{n_0 m}{n} \tag{2.20}$$

and define

$$\omega_\alpha(S) = \begin{cases} 1 \text{ if } |S \cap J_\alpha| \geq K_0 \\ 0 \text{ otherwise.} \end{cases} \tag{2.21}$$

Hence $\sum_\alpha \omega_\alpha(S) \leq \frac{|S|}{K_0}$ and by (2.20)

$$\sum_\alpha \Big[\sum_{|S| < n_0} |\hat{\Lambda}(S)|^2 \big(1 - \omega_\alpha(S)\big) \Big] \geq \frac{n}{2m} \sum_{|S| < n_0} |\hat{\Lambda}(S)|^2.$$

We may therefore redefine for some $\alpha = 1, \dots, \left[\frac{n}{m}\right]$ the function f as

$$f = f_\alpha = \sum_{\substack{|S| \le n_0 \\ |S \cap J_\alpha| < K_0}} \hat{f}(S) w_S \tag{2.22}$$

and still satisfy

$$\sum \hat{\Lambda}(S) \hat{f}(S) \gtrsim \sum_{|S| < n_0} |\hat{\Lambda}(S)|^2. \tag{2.23}$$

Note that $\|f\|_2 \le \|\Lambda\|_2 \le \sqrt{Nn}$, $\|f\|_1 \le \sqrt{n} N$ and, cf. [B1], [B2]

$$\|\hat{f}\|_1 < (Cn)^{n_0}. \tag{2.24}$$

Recalling (2.16), (2.17), it follows that

$$(2.6) + (2.7) + (2.9) < (Cn)^{n_0} N \Big(\frac{1}{\sqrt{B}} + \frac{B}{T} \Big) < (Cn)^{n_0} \frac{N}{\sqrt{B}} \tag{2.25}$$

provided $B < \sqrt{T}$.

Obtaining the required bounds on

$$\sum_{\substack{q < B \\ q\, sf \text{ and odd}}} \frac{\kappa(q)}{q} \tag{2.26}$$

and

$$\sum_{\substack{q_3 < B \\ q_3\, sf \text{ and odd}}} \frac{\alpha(q_1, q_3)}{q_3} \tag{2.27}$$

in (2.8), (2.10) is a non-trivial and crucial input in [B1], [B2]. Essential use is made in the argument of the property $|S \cap J_\alpha| < K_0$ in (2.22). The basic idea is as follows. Assume that $|S \cap J_\alpha| < K_0$ if $\hat{f}(S) \ne 0$. Then the function f restricted to a translate of a progression of the form $2^j.I$, $I = \{1, \dots, m\}$, has a low order Walsh expansion.

Switching to an expansion in the trigonometric system (with suitable approximation of the Walsh functions) permits then to analyze the l.h.s. in (2.14), (2.15).

This analysis was carried out in [B1], §4, 5 for a function f of the form $1_{[x_{j_1} = \alpha_1, \dots, x_{j_r} = \alpha_r]}$. Some adjustments of this rather tedious argument (noting also that $\mathbb{E}[f] = 0$ here, by assumption) permit us to obtain an estimate of the form

$$(2.26), (2.27) < e^{-c \frac{n}{n_0}}. \tag{2.28}$$

Note that in [B2], §4, this estimate on (2.27) is proven for a function f of the type considered above.

From (2.25), (2.29), it follows that

$$\Big|\sum_1^N \Lambda(x) f(x)\Big| < N\big((Cn)^{n_0} B^{-\frac12} + e^{-c\frac{n}{n_0}}(n^3 + |\mathcal{B}|)\big) \tag{2.29}$$

where the size of $\mathcal{B}$ is controlled from classical zero-density estimates

$$|\mathcal{B}| \lesssim (TB)^{8\eta_*}, \eta_* = C\frac{\log T}{n} \tag{2.30}$$

(cf. [B1], §6). Taking, according to (2.16)

$$\log B = c \min\Big(\frac{n}{\sqrt{n_0}}, \frac{n^{4/7}}{(\log n)^{3/7}}\Big) \tag{2.31}$$

and recalling the assumption on n_0 in Proposition 1, gives

$$(2.29) < \exp\Big\{-c\min\Big[\frac{n}{n_0}, \frac{n^{4/7}}{(\log n)^{3/7}}\Big]\Big\}. \tag{2.32}$$

Inequality (2.4) follows.

3. The function $\tilde\Lambda$ and proof of the theorem

Recall that $\tilde\Lambda$ is defined as

$$\tilde\Lambda = \sum_x \Lambda(x) \sum_{j|x_j=1} \delta_{x\backslash\{j\}} \tag{3.1}$$

with $\delta_{x\backslash\{j\}}$ the Dirac at $x\backslash\{j\}$. From (1.5), implied by monotonicity,

$$\langle\tilde\Lambda, f\rangle \leq n\langle\Lambda, f\rangle. \tag{3.2}$$

We start by evaluating $\|\tilde\Lambda\|_2$. Clearly, by (3.1)

$$\begin{aligned}\sum_{y\leq N}\tilde\Lambda(y)^2 &\leq \sum_{y\leq N}\sum_{0\leq j,k<n}\Lambda(y+2^j)\Lambda(y+2^k)\\ &= \sum_{0\leq j,k<n}\sum_{x<2N}\Lambda(x)\Lambda(x+2^k-2^j).\end{aligned} \tag{3.3}$$

Using a simple upper bound sieve (cf. [I-K], Th. 6.7), one has for $j\neq k$

$$\sum_{x<2N}\Lambda(x)\Lambda(x+2^k-2^j) \lesssim N \tag{3.4}$$

implying that

$$\|\tilde\Lambda\|_2 \lesssim n\sqrt{N}. \tag{3.5}$$

Next, we compute the F-W coefficient of $\tilde\Lambda$.

Rewrite (3.1) as

$$\tilde{\Lambda} = \sum_x \Lambda(x) \Big[\delta_{(0,x_1,\ldots,x_{n-1})} + \delta_{(x_0,0,x_2,\ldots,x_{n-1})} + \cdots + \delta_{(x_0,\ldots,x_{n-2},0)} - \Big(n - \sum x_j\Big) \delta_x \Big].$$

For $S \subset \{0, 1, \ldots, n-1\}$, we get

$$\sum_x \Lambda(x) w_S(0, x_1, \ldots, x_{n-1}) = \begin{cases} N\hat{\Lambda}(S) & \text{if } 0 \notin S \\ N\hat{\Lambda}(S\backslash\{0\}) & \text{if } 0 \in S. \end{cases} \tag{3.6}$$

Also from the formulas $\varepsilon_j = 1 - 2x_j$, $n - \sum\limits_0^{n-1} x_j = \sum\limits_j^{n-1} \frac{1+\varepsilon_j}{2}$, we obtain

$$\begin{aligned} &\sum_x \Lambda(x)(n - \sum x_j) w_S(x) = N\langle \Lambda, \Big(\sum_0^{n-1} \frac{1+\varepsilon_j}{2} \Big) \prod_{j\in S} \varepsilon_j \rangle = \\ &N\Big\{ \frac{n}{2}\hat{\Lambda}(S) + \frac{1}{2} \sum_{j\in S} \hat{\Lambda}(S\backslash\{j\}) + \frac{1}{2} \sum_{j \notin S} \hat{\Lambda}(S \cup \{j\}) \Big\}. \end{aligned} \tag{3.7}$$

Hence

$$\hat{\tilde{\Lambda}}(S) = \Big(\frac{n}{2} - |S|\Big) = \tilde{\Lambda}(S) + \frac{1}{2} \sum_{j\in S} \hat{\Lambda}(S\backslash\{j\}) - \frac{1}{2} \sum_{j\notin S} \hat{\Lambda}(S \cup \{j\}). \tag{3.8}$$

In particular, it follows that

$$\mathbb{E}[\tilde{\Lambda}] = \frac{n}{2} - \frac{1}{2} \sum_j \hat{\Lambda}(\{j\}) = \frac{n-1}{2} + o(1) \tag{3.9}$$

$$\|\tilde{\Lambda}\|_1 \sim nN. \tag{3.10}$$

Combined with (3.5), it follows that the distribution $\tilde{\Lambda}$ is essentially 'flat'.

Write using F-W expansion and (3.9)

$$\langle f, \tilde{\Lambda} \rangle = \mathbb{E}[f]\, \mathbb{E}[\tilde{\Lambda}] + (3.12) + (3.13) > \Big(\frac{n}{2} - 1\Big) \mathbb{E}[f] + (3.12) + (3.13) = (3.11) \tag{3.11}$$

with

$$(3.12) = \sum_{\substack{S \neq \phi \\ |S| < K\sqrt{n}}} \hat{f}(S) \hat{\tilde{\Lambda}}(S)$$

and

$$(3.13) = \sum_{|S| \geq K\sqrt{n}} \hat{f}(S) \hat{\tilde{\Lambda}}(S)$$

(K a large parameter).

In view of (1.3), (3.5)

$$(3.13) \lesssim K^{-\frac{1}{2}} \frac{\|\tilde{\Lambda}\|_2}{\sqrt{N}} \lesssim K^{-\frac{1}{2}} n. \tag{3.14}$$

It follows from (3.8) and Proposition 2 that certainly

$$\sum_{2<|S|\leq K\sqrt{n}} |\hat{\tilde{\Lambda}}(S)|^2 < n^2 e^{-n^{3/7}} \tag{3.15}$$

assuming

$$K < n^{\frac{1}{14}} (\log n)^{-2}. \tag{3.16}$$

For $S = \{0\}$, by (3.8), (3.15)

$$\begin{aligned}\hat{\tilde{\Lambda}}(\{0\}) &= \Big(\frac{n}{2} - 1\Big)\hat{\Lambda}(\{0\}) + \frac{1}{2}\hat{\Lambda}(\phi) + O(e^{-\frac{1}{3}n^{3/7}}) \\ &= \frac{3-n}{2} + O\big(e^{-\frac{1}{3}^{n^{3/7}}}\big)\end{aligned}$$

since $\hat{\Lambda}(\{0\}) = \mathbb{E}[\Lambda\varepsilon_0] = -\hat{\Lambda}(\phi)$ by parity.

If $S = \{j\}, 0 < j < n$,

$$\hat{\tilde{\Lambda}}(\{j\}) = \frac{1}{2}\hat{\tilde{\Lambda}}(\phi) + O(e^{-\frac{1}{3}n^{3/7}}) = \frac{1}{2} + O(e^{-\frac{1}{3}n^{3/7}}).$$

For $S = \{0, j\}$,

$$\hat{\tilde{\Lambda}}(S) = \frac{1}{2}\hat{\Lambda}(\{0\}) + O(e^{-\frac{1}{3}n^{3/7}}) = -\frac{1}{2} + O(e^{-\frac{1}{3}n^{3/7}})$$

and $|\hat{\tilde{\Lambda}}(S)| < O(e^{-\frac{1}{3}n^{3/7}})$ if $|S| = 2, 0 \notin S$.

Therefore

$$\begin{aligned}(3.12) &= -\frac{n-3}{2}\hat{f}(\{0\}) + \frac{1}{2}\sum_{0<j<n} \hat{f}(\{j\}) - \frac{1}{2}\sum_{0<j<n} \hat{f}(\{0,j\}) + O(e^{-\frac{1}{3}n^{3/7}}) \\ &= -\frac{n-3}{2}\mathbb{E}[f.\varepsilon_0] + \frac{1}{2}\sum_{0<j<n} \mathbb{E}[(1-\varepsilon_0)\varepsilon_j f] + O(e^{-\frac{1}{3}n^{3/7}}) \\ &= \frac{n-3}{2}\mathbb{E}[f] + \sum_{0<j<n} \mathbb{E}[\varepsilon_j f] + O(e^{-\frac{1}{3}n^{3/7}}) \\ &= \frac{n-3}{2}\mathbb{E}[f] + O(\sqrt{n})\end{aligned} \tag{3.17}$$

again using that supp $f \subset [x_0 = 1]$.

Substituting (3.14), (3.17) into (3.11),

$$\langle f, \tilde{\Lambda}\rangle > \big(n - O(1)\big)\mathbb{E}[f] + O(\sqrt{n}) + O(K^{-\frac{1}{2}}n). \tag{3.18}$$

From (3.16), (3.2)

$$\langle f, \Lambda\rangle \geq \mathbb{E}[f] + O(n^{-\frac{1}{28}}\log n). \tag{3.19}$$

This gives (1.1), with in fact $n^{-\frac{1}{28}}\log n$ in place of $o(1)$.

References

[B1] J. Bourgain, *Prescribing the binary digits of the primes*, to appear in Israel J. Math.

[B2] J. Bourgain, *On the Fourier-Walsh spectrum of the Moebius function*, to appear in Israel J. Math.

[B3] J. Bourgain, *A prime number theorem for the majority function*, preprint.

[B-T] N. Bshouty, C. Tamon, *On the Fourier spectrum of monotone functions*, Journal of the ACM, Vol. 43, N4 (1996), 747–770.

[Gr] B. Green, *On (not) computing the Moebius function using bounded depth circuits*, ArXiv: 1103,4991.

[Ka] G. Kalai, *Private communications*

[I-K] H. Iwaniec, E. Kowalski, Analytic Number theory, AMS.

School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540.

E-mail address: `bourgainmath.ias.edu`